



(5') 1 SDLEQERRAKEKLQEQQ
18 SDLEQDRLAKEKLQEQQ
35 SDLEQERLAKEKLQEQQ
52 SDLEQERRAKEKLQEQQ
69 SDLEQERRAKEKLQEQQ
86 SDLEQDRLAKEKLQEQQ
103 SDLEQERRAKEKLQEQQ
120 SDLEQERRAKEKLQEQQ
137 SDLEQERLAKEKLQEQQ
154 SDLEQERRAKEKLQEQQ
171 SDLEQERRAKEKLQEQQ
188 SDLEQERRAKEKLQEQQ
205 RDLEQ

210 RKADTKKNLERKKEHGDILAEDLYGRLEIP
240 AIELPSENERGYYPHQSSLPQDNRGNSRD
270 SKEISIIIEKTNRESITTNVEGRRDIHKGHL
300 EEKDGSIKPEQKEDKS 316 (3') (SEQ ID NO:31)

FIGURE 1

(5') 1 AAAGCGATCTAGAACAAAGAGAGACGTGCTAAAGAAAAAGTTGCAAGAACAAAC
52 AAAGCGATTTAGAACAAAGATAGACTTGCTAAAGAAAAAGTTACAAGAGCAGC
103 AAAGCGATTTAGAACAAAGAGAGACCTTGCTAAAGAAAAAGTTGCAAGAACAAAC
154 AAAGCGATCTAGAACAAAGAGAGACGTGCTAAAGAAAAAGTTGCAAGAACAAAC
205 AAAGCGATTTAGAACAAAGAGAGACGTGCTAAAGAAAAAGTTGCAAGAACAAAC
256 AAAGCGATTTAGAACAAAGATAGACTTGCTAAAGAAAAAGTTACAAGAGCAGC
307 AAAGCGATTTAGAACAAAGAGAGACGTGCTAAAGAAAAAGTTGCAAGAACAAAC
358 AAAGCGATTTAGAACAAAGAGAGACGTGCTAAAGAAAAAGTTGCAAGAACAAAC
409 AAAGCGATTTAGAACAAAGAGAGACCTTGCTAAAGAAAAAGTTGCAAGAACAAAC
460 AAAGCGATTTAGAACAAAGAGAGACGTGCTAAAGAAAAAGTTGCAAGAACAAAC
511 AAAGCGATTTAGAACAAAGAGAGACGTGCTAAAGAAAAAGTTGCAAGAACAAAC
562 AAAGCGATTTAGAACAAAGAGAGACGTGCTAAAGAAAAAGTTGCAAGAGCAGC
613 AAAGAGATTTAGAACAA
630 AGGAAGGCTGATACGAAAAAAATTTAGAAAGAAAAAGGAACATGGAGAT
681 ATATTAGCAGAGGATTTATATGTCGTTTAGAAATACCACTATAGAACTT
732 CCATCAGAAAAATGAACGTGGATATTATATACCAATCATCTCTTTACCT
783 CAGGACAAACAGAGGGAATAGTAGAGATTCCAAGGAAATATCTATAATAGAA
834 AAACAAATAGAGAACTCTATTACAACAAATGTTGAAGGACGAAGGGATATA
885 CATAAAGGACATCTTGAAGAAAAAGAAAGATGGTTCAATAAAACCAGAACAA
936 AAAGAAGATAAATCT 950 (3') (SEQ ID NO: 32)

FIGURE 2

^E
RDELFNELLNSVDVNGEVKENILEESQVND~~Q~~DIFNSLVKSVQQEQQ
HNVEEKVEESVEENDEESVEENVEENVEENDDGSVASSVEESI
ASSVDESIDSSIEENVAPTVEEIVAPTVEEIVAPSVVEKCAPSVE
ESVAPSVEESVAEMLKER (SEQ ID NO: 24)

FIGURE 3

5' GAA TTC CGT GAT GAA CTT TTT AAT GAA TTA TTA AAT AGT GTA GAT
 GTT AAT GGA GAA GTA AAA GAA AAT ATT TTG GAG GAA AGT CAA GTT AAT
 GAG GAT ATT TTT AAT AGT TTA GTA AAA AGT GTT CAA CAA GAA CAA CAA
 CAC AAT GTT GAA GAA AA AGT TGA AGA AAG TGT AGA AGA AA ATG ACG
 AAG AAA GTG TAG AAG AAA ATG TAG AAG AAA ATG TAG AAG AAA ATG
 ACG ACG GAA GTG TAG CCT CAA GTG TTG AAG AAA GTA TAG CTT CAA GTG
 TTG ATG AAA GTA TAG ATT CAA GTA TTG AAG AAA ATG TAG CTC CAA CTG
 TTG AAG AAA TCG TAG CTC CAA CTG TTG AAG AAA TTG TAG CTC CAA GTG
 TTG TAG AAA AGT GTG CTC CAA GTG TTG AAG AAA GTG TAG CTC CAA GTG
 TTG AAG AAA GTG TAG CTG AAA TGT TGA AGG AAA GGA ATT C 3' (SEQ ID No: 33)

FIGURE 4

<u>DELFNELLNSVDVNGEVKENILEESQ</u>	<u>NSRDSKEISIIIEKTNRESIITINVEGRDIIHK</u>	<u>LSA-TER (see ID NO: 23)</u>
<u>LEESQVNDDDIFNSLVKSVQEQQHNV</u>		<u>729S-NRI (see ID NO: 26)</u>
<u>VEKCAPSVEESVAPSVEESVAEMLKER</u>		<u>729S-NRII (see ID NO: 27)</u>
		<u>729S-Rep (see ID NO: 28)</u>

FIGURE 5

NUCLEOTIDE SEQUENCE OF THE LSA GENE
5' END

(NON-CODING 5' END)

1 AAAGTATACATCTTCCTTCTTTACTTCTTAAA

(CODING 5' END, UNIQUE)

33 ATGAAACATATTTTGTACATATCATTTTACTTTATCCTTGTTAATTTATTG
84 ATATTTTCATATAAATGGAAAGATAATAAAGAATTCTGAAAAAGATGAAATCA
135 TAAAATCTAACTTGAGAAGTGTTCTTCAAATTCTAGGAATCGAATAAATGA
186 GGAAATCACGAGAAGAAACACGTTTTATCTCATAATTCATATGAGAAAAC
237 AAAAATAATGAAAATAATAAATTTTTTCGATAAGGATAAAGAGTTAACGATGT
288 CTAATGTAAAAAATGTGTCACAAACAAATTTCAAAGTCTTTTAAGAAATCT
339 TGGTGTTTCAGAGAATATATTCCTTAAAGAAAATAAATTAAATAAGGAAGGG
390 AAATTAATTGAACACATAATAAATGATGATGACGATAAAAAAAAAAATATATTA
441 AAGGGCAAGACGAAAACAGACAAGAAGATCTTGAAGAAAAAGCA

(CODING 5' END, repetitive)

492 GCTAAAGAAAAGTTACAGGGGCAACAAAGCGATTGAGAACAAGAGAGACGT
543 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTGAGAACAAGAGAGACTT
594 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTGAGAACAAGAGAGACGT
645 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTGAGAACAAGAGAGACTT
696 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTGAGAACAAGAGAGACGT
747 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTGAGAACAAGAGAGACGT
798 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTGAGAACAAGAGAGACTT
849 GCTAAAGAAAAGTTACAAGAGCAGCAAAGCGATTGAGAACAAGATAGACTT
900 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTGAGAACAAGAGAGACGT
951 GCTAAAGAAAGGTTGCAAGAACAACAAAGCGATTGAGA 988 (SEQ ID NO: 34)

FIGURE 6

(SEQ ID NO: 35)

(SEQ ID NO: 36)

(SEQ ID NO: 36)

FIGURE 7A

NUCLEOTIDE SEQUENCE OF THE LSA GENE
3' END

(CODING 3' END, REPETITIVE)

1 CAAGAACAACAAAGCGATCTAGAACAAGAGAGACGT
37 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGATAGACTT
88 GCTAAAGAAAAGTTACAAGAGCAGCAAAGCGATTTAGAACAAGAGAGACTT
139 GCTAAGAAAAGTTGCAAGAACAACAAAGCGATCTAGAACAAGAGAGACGT
190 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
241 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGATAGACTT
292 GCTAAAGAAAAGTTACAAGAGCAGCAAAGCGATTTAGAACAAGAGAGACGT
343 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
394 GCTAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACTT
445 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
496 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
547 GCTAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
598 GCTAAAGAAAAGTTGCAAGAGCAGCAAAGAGATTTAGAACAA

(CODING 3' END, UNIQUE)

640 AGGAAGGCTGATACGAAAAAAATTTAGAAAGAAAAAAGGAACATGGAGAT
691 ATATTAGCAGAGGATTTATATGGTCGTTTAGAAATACCAGCTATAGAACTT
742 CCATCAGAAAATGAACGTGGATATTATATACCACATCAATCTTCTTTACCT
793 CAGGACAACAGAGGGAATAGTAGAGATTCCAAGGAAATATCTATAATAGAA
844 AAAACAAATAGAGAATCTATTACAACAAATGTTGAAGGACGAAGGGATATA
895 CATAAAGGACATCTTGAAGAAAAGAAAGATGGTTCAATAAAACCAGAACAA
946 AAAGAAGATAAATCTGCTGACATACAAAATCATACATTAGAGACAGTAAAT
997 ATTTCTGATGTTAATGATTTTCAAATAAGTAAGTATGAGGATGAAATAAGT
1048 GCTGAATATGACGATTCATTAATAGATGAAGAAGAAGATGATGAAGACT
1099 TAGACGAATTTAAGCCTATTGTGCAATATGACAATTTCCAAGATGAAGAAA
1150 ACATAGGAATTTATAAAGAACTAGAAGATTTGATAGAGAAAAATGAAAATT
1201 TAGATGATTTAGATGAAGGAATAGAAAAATCATCAGAAGAATTATCTGAAG
1252 AAAAAATAAAAAAAGGAAAGAAATATGAAAAACAAAGGATAATAATTTTA
1303 AACCAAATGATAAAAGTTTGTATGATGAGCATATTAAAAAATATAAAAATG
1354 ATAAGCAGGTTAATAAGGAAAAGGAAAAATTCATAAAATCATTGTTTCATA
1405 TATTTGACGGAGACAATGAAATTTTACAGATCGTGGATGAGTTATCTGAAG
1456 ATATACTAAATATTTTATGAACTATAA (stop) (SEQ ID NO: 39)

(NON-CODING 3' END)

1485 AAGGTTATATATTT 1498

FIGURE 8

LSA.3'.ALL -> 1-phase Translation

DNA sequence 1496 b.p. (SEQ ID NO: 40) (SEQ ID NO: 41) linear
CAAGAACAACAA ... GGTATATATTT

1 / 1	31 / 11	
(SEQ ID NO: 42) CAA GAA CAA CAA AGC GAT CTA GAA CAA GAG AGA CGT GCT AAA GAA AAG TTG CAA GAA CAA		
(SEQ ID NO: 43) gln gln gln ser asp leu gln gln gln ala lys gln lys leu gln gln gln		
61 / 21	91 / 31	
CAA AGC GAT TTA GAA CAA GAT AGA CTT GCT AAA GAA AAG TTA CAA GAG CAG CAA AGC GAT		
gln ser asp leu gln gln asp arg leu ala lys gln lys leu gln gln gln ser asp		
121 / 41	151 / 51	
TTA GAA CAA GAG AGA CTT GCT AAA GAA AAG TTG CAA GAA CAA AGC GAT CTA GAA CAA		
leu gln gln gln arg leu ala lys gln lys leu gln gln gln ser asp leu gln gln		
181 / 61	211 / 71	
GAG AGA CGT GCT AAA GAA AAG TTG CAA GAA CAA AGC GAT TTA GAA CAA GAG AGA CGT		
glu arg arg ala lys gln lys leu gln gln ser asp leu gln gln gln arg arg		
241 / 81	271 / 91	
GCT AAA GAA AAG TTG CAA GAA CAA AGC GAT TTA GAA CAA GAT AGA CTT GCT AAA GAA		
ala lys gln lys leu gln gln gln ser asp leu gln gln asp arg leu ala lys gln		
301 / 101	331 / 111	
AAG TTA CAA GAG CAG CAA AGC GAT TTA GAA CAA GAG AGA CGT GCT AAA GAA AAG TTG CAA		
lys leu gln gln gln ser asp leu gln gln arg arg ala lys gln lys leu gln		

FIGURE 9A

1261 / 421 1291 / 431
 AAA GGA AAG AAA TAT GAA AAA ACA AAG GAT AAT AAT TTT AAA CCA AAT GAT AAA AGT TTG
 lys gly lys lys tyr glu lys thr lys asp asn phe lys pro asn asp lys ser leu
 1321 / 441 1351 / 451
 TAT GAT GAG CAT ATT AAA TAT AAA AAT GAT AAG CAG GTT AAT AAG GAA AAG GAA AAA
 tyr asp glu his ile lys lys tyr lys asn asp lys gln val asn lys glu lys glu lys
 1381 / 461 1411 / 471
 TTC ATA AAA TCA TTG TTT CAT ATA TTT GAC GAC AAT GAA ATT TTA CAG ATC GTG GAT
 phe ile lys ser leu phe his ile phe asp gly asp asn glu ile leu gln ile val asp
 1441 / 481 1471 / 491
 GAG TTA TCT GAA GAT ATA ACT AAA TAT TTT ATG AAA CTA TAA AAG GTT ATA TAT
 glu leu ser glu asp ile thr lys tyr phe met lys leu ~~COCH lys val ile tyr~~

* Strike from Fig.

FIGURE 9D

LSN.3'STOP -> 1-phase Translation

DNA sequence 1482 b.p. CAAGAACAACAA ... ATGAAACTATAA linear

(seq ID No: 44) (seq ID No: 45)

1 / 1	31 / 11	
(seq ID No: 44) CAA GAA CAA CAA AGC GAT CTA GAA CAA GAG AGA CGT GCT AAA GAA AAG TTG CAA GAA GAA		
(seq ID No: 47) gln glu gln gln ser asp leu glu gln gln glu arg ala lys glu lys leu gln glu gln		
61 / 21	91 / 31	
CAA AGC GAT TTA GAA CAA GAT AGA CTT GCT AAA GAA AAG TTA CAA GAG CAG CAA AGC GAT		
gln ser asp leu glu gln asp arg leu ala lys glu lys leu gln glu gln ser asp		
121 / 41	151 / 51	
TTA GAA CAA GAG AGA CTT GCT AAA GAA AAG TTG CAA GAA CAA AGC GAT CTA GAA CAA		
leu glu gln glu arg leu ala lys glu lys leu gln glu gln ser asp leu glu gln		
181 / 61	211 / 71	
GAG AGA CGT GCT AAA GAA AAG TTG CAA GAA CAA CAA AGC GAT TTA GAA CAA GAG AGA CGT		
glu arg arg ala lys glu lys leu gln glu gln ser asp leu glu gln glu arg arg		
241 / 81	271 / 91	
GCT AAA GAA AAG TTG CAA GAA CAA CAA AGC GAT TTA GAA CAA GAT AGA CTT GCT AAA GAA		
ala lys glu lys leu gln glu gln ser asp leu glu gln asp arg leu ala lys glu		
301 / 101	331 / 111	
AAG TTA CAA GAG CAG CAA AGC GAT TTA GAA CAA GAG AGA CGT GCT AAA GAA AAG TTG CAA		
lys leu gln glu gln ser asp leu glu gln glu arg ala lys glu lys leu gln		

FIGURE 10A

1321	/	441		1351	/	451	
TAT	GAT	GAG	CAT	ATT	AAA	TAT	AAA
tyr	asp	glu	his	ile	lys	tyr	lys
1381	/	461		1411	/	471	
TTC	ATA	AAA	TCA	TTG	TTT	GAC	ATA
phe	ile	lys	ser	leu	phe	ile	lys
1441	/	481		1471	/	491	
GAG	TTA	TCT	GAA	GAT	ATA	ACT	AAA
glu	leu	ser	glu	asp	ile	thr	lys
						tyr	phe
						met	lys
						leu	GAG
						lys	val
						ile	tyr
						tyr	tyr

*DELETE FROM FIGURE

FIGURE 10D